

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-191225

(43) 公開日 平成8年(1996)7月23日

(51) Int.Cl.⁶

識別記号

庁内整理番号

F I

技術表示箇所

H 0 3 G 5/24

G 1 0 K 15/00

H 0 4 S 1/00

7/00

L

F

G 1 0 K 15/ 00

M

審査請求 未請求 請求項の数 5 F D (全 10 頁)

(21) 出願番号

特願平7-18576

(22) 出願日

平成7年(1995)1月9日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 小川 理子

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 田村 忠司

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72) 発明者 片山 崇

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74) 代理人 弁理士 岡本 宜喜

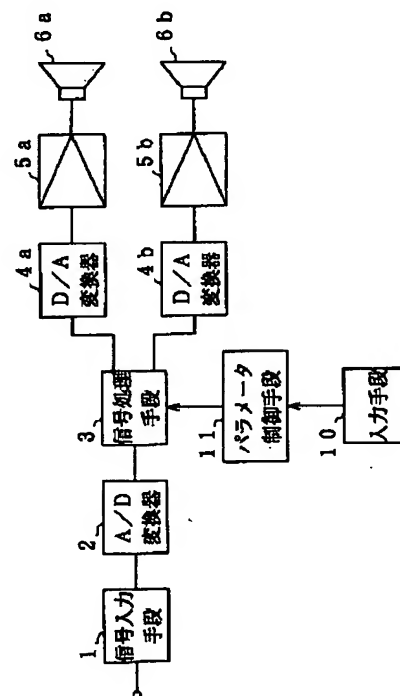
最終頁に続く

(54) 【発明の名称】 音場再生装置

(57) 【要約】

【目的】 ヘッドホンやインナホンを使用してあらゆる方向の自然な音像定位感、距離感や広がり感を得られるようにすること。

【構成】 信号入力手段1の音声信号をA/D変換器2に与え、サンプリング周波数 f_s でデジタル信号に変換する。信号処理手段3はデジタル信号を入力すると、FIRフィルタを用いて音像定位のための演算を行う。このとき、畳込み係数のオーバーサンプリングをし、受聴者の両耳間隔差に応じて畳込み係数を乗じるタップ位置を細かく制御する。こうして受聴者の頭部形状や大きさ、受聴者と音像との位置関係に応じて畳込み係数を制御する。こうして演算された音声信号をD/A変換器4に与え、アンプ5を介してヘッドホンのスピーカ6a、6bに出力する。



【特許請求の範囲】

【請求項 1】 入力音声信号をデジタル信号に変換する A/D 変換器と、前記 A/D 変換器により変換されたデジタル信号が入力され、受聴者の頭部形状や大きさ、受聴者と音像の位置及び距離に応じた音響信号となるように信号処理を行う信号処理手段と、受聴者の頭部形状や大きさ、及び受聴者と音像の位置を入力する入力手段と、前記入力手段により入力された内容に応じた音響特性が得られるようにパラメータを設定し、前記信号処理手段に前記パラメータを与えるパラメータ制御手段と、前記信号処理手段から出力される信号をアナログ信号に変換する D/A 変換器と、前記 D/A 変換器の出力を増幅するアンプと、前記アンプの出力を音響変換する左右の音響変換器と、を具備する音場再生装置であって、前記信号処理手段は、前記入力手段により入力された受聴者と音像の位置関係に応じて、畳込み演算により音像を定位させる音像定位手段と、前記音像定位手段を介して前記パラメータ制御手段から基準の畳込み係数をパラメータとして入力し、これらのパラメータを時系列の離散化信号に変換し、前記 A/D 変換器のサンプリング周波数以上の値で前記パラメータをオーバーサンブルするオーバーサンブル手段と、前記オーバーサンブル手段で処理されたパラメータに対して、所定の周波数特性以下の成分を通過させるフィルタ手段と、前記フィルタ手段から出力されるパラメータを、前記オーバーサンブル手段のサンプリング周波数以下の値でダウンサンブルし、得られた時系列の信号を夫々の畳込み係数に並列変換し、その値を補正された畳込み係数として前記音像定位手段に与えるダウンサンブル手段と、を有するものであることを特徴とする音場再生装置。

【請求項 2】 前記フィルタ手段は、前記 A/D 変換器のサンプリング周波数の $1/2$ 以下の周波数成分を通過させるものであることを特徴とする請求項 1 記載の音場再生装置。

【請求項 3】 前記入力手段は、受聴者の頭部形状や大きさ、及び受聴者と音像の位置を入力すると共に、音場の広がりも同時に入力するものであり、前記信号処理手段の出力部に、前記入力手段より得られる音場の広さに対応した残響音を加える残響音作成手段を設けたことを特徴とする請求項 1 記載の音場再生装置。

【請求項 4】 前記パラメータ制御手段は、音響空間の音源から壁で反射して受聴者に入射される反射音において、受聴者から見た音源と反射音方向と角度が入力されると、前記角度を音響空間における受聴位置

又は音像位置に変換してパラメータを前記信号処理手段に与えるものであることを特徴とする請求項 1 記載の音場再生装置。

【請求項 5】 前記入力手段は、外部から音場に関する制御信号を受信し、前記パラメータ制御手段に与えるパラメータ受信手段を含むものであることを特徴とする請求項 1 記載の音場再生装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、AV（オーディオ・ビジュアル）機器において、任意の受聴者に対して自然な定位感、臨場感のある音響再生を行うために、個人の両耳間隔、頭部形状を考慮した音場再生装置に関するものである。

【0002】

【従来の技術】近年、映像・音響を中心としたマルチメディア分野においては、映像の 3 次元コンピュータグラフィック化に伴い、3 次元の立体的な音響再生に対応するハードウェアの開発が望まれている。また、TV やパソコンなどを対象とした家庭用ゲーム機のようなインタラクティブな機器が普及しつつある。これに対応すべく 2 チャンネルステレオスピーカのみならず、イヤースピーカ、ヘッドホン、インナホンをを用いた音場再生において、音の定位感、距離感、広がり感をインタラクティブに制御する開発も行われている。

【0003】図 11 は、スピーカを用いた従来の音場再生装置の一例を示すブロック図である。同図において、信号入力手段 101 はオーディオ信号を入力する入力手段であって、その出力はフィルタ 102 を介してアンプ 103a、103b に与えられる。フィルタ 102 は 2 チャンネル信号の周波数特性、音量、時間遅延を制御するものであり、音像位置入力手段 104 により制御される。音像位置入力手段 104 は受聴者の音の定位感を制御するために音像位置を入力する手段である。アンプ 103a、103b で増幅された音声信号はスピーカ 105a、105b に与えられる。

【0004】このように構成された従来の音場再生装置の動作について説明する。信号入力手段 101 により入力された音声信号は、音像位置入力手段 104 で入力された音現の位置に応じた周波数特性、音量、時間遅延でスピーカ 105a、105b から音声を再生するようフィルタ 102 を制御する。右耳より左耳の方が時間遅延が大きく、中高域の周波数特性で音圧レベルが低くなると、一般に、音像が右方にあると感じる。フィルタ 102 はこのような制御を行い、受聴者の音像定位感を制御する。フィルタ 102 により制御された信号はアンプ 103a、103b により増幅され、スピーカ 105a、105b により再生される。このような信号処理を行うことにより、音の定位感を制御することができる。

【0005】

【発明が解決しようとする課題】しかしながら上記のような構成では、再生音を聴いても、受聴者の頭部、耳部の形状の相違による微妙な特性の変化を制御することは困難であった。特にヘッドホンや外耳道に挿入されたインナホンの場合は、音像における前方の頭外感、定位感、距離感を実現することができなかった。

【0006】本発明は、このような従来の問題点に鑑みてなされたものであって、任意の受聴者に対して、あらゆる方向の自然な音像定位感、距離感が得られるような、音場再生装置を提供することを目的とするものである。

【0007】

【課題を解決するための手段】本発明は入力音声信号をデジタル信号に変換するA/D変換器と、A/D変換器により変換されたデジタル信号が入力され、受聴者の頭部形状や大きさ、受聴者と音像の位置及び距離に応じた音響信号となるように信号処理を行う信号処理手段と、受聴者の頭部形状や大きさ、及び受聴者と音像の位置を入力する入力手段と、入力手段により入力された内容に応じた音響特性が得られるようにパラメータを設定し、信号処理手段にパラメータを与えるパラメータ制御手段と、信号処理手段から出力される信号をアナログ信号に変換するD/A変換器と、D/A変換器の出力を増幅するアンプと、アンプの出力を音響変換する左右の音響変換器と、を具備する音場再生装置であって、信号処理手段は、入力手段により入力された受聴者と音像の位置関係に応じて、畳込み演算により音像を定位させる音像定位手段と、音像定位手段を介してパラメータ制御手段から基準の畳込み係数をパラメータとして入力し、これらのパラメータを時系列の離散化信号に変換し、A/D変換器のサンプリング周波数以上の値でパラメータをオーバーサンプリングするオーバーサンプリング手段と、オーバーサンプリング手段で処理されたパラメータに対して、所定の周波数特性以下の成分を通過させるフィルタ手段と、フィルタ手段から出力されるパラメータを、オーバーサンプリング手段のサンプリング周波数以下の値でダウンサンプルし、得られた時系列の信号を夫々の畳込み係数に並列変換し、その値を補正された畳込み係数として音像定位手段に与えるダウンサンプル手段と、を有することを特徴とするものである。

【0008】

【作用】このような特徴を有する本発明によれば、入力手段によって受聴者の頭部形状や大きさ、受聴者と音像の位置や距離を入力設定する。音声信号が信号入力手段を介して入力されると、A/D変換器によりデジタル信号に変換され、信号処理手段に入力される。パラメータ制御手段には、入力手段によって設定された内容に応じた音響特性が得られるようにパラメータを選択する。そして制定条件に合致するパラメータを畳込み係数として信号処理手段に与える。信号処理回路で演算された信号は

D/A変換器によってアナログ信号に変換され、アンプを介して音響変換器より音出力される。

【0009】特に信号処理手段では、入力手段により入力された受聴者の頭部形状や大きさ、受聴者と音像の位置関係に応じて音像定位させ、その出力をオーバーサンプルし、所定の周波数特性でフィルタリングしたのち、元のサンプリング周波数に近い値でダウンサンプルして音響信号を出力するようにしている。このとき、任意の受聴者の微妙な両耳間隔差を出力信号に反映させるために、必要なサンプリング周波数でオーバーサンプルする。また、オーバーサンプリングされた信号がフィルタ手段に与えられると、A/D変換器のサンプリング周波数の1/2以下の成分のものが通過する。こうするとダウンサンプリング手段でダウンサンプルされる際の歪は少なくなる。このように処理すると任意の受聴者に対して微妙な両耳間隔差に対応し、あらゆる方向に対して自然な音像定位と音場再生が得られる。

【0010】

【実施例】本発明の実施例における音場再生装置について図面を参照しつつ説明する。まず図3の音像定位の原理図を用いて、ヘッドホンスピーカ6a、6bから受聴者7の左後方に音を定位させる方法について説明する。本図において信号入力手段1は音声信号S(t)を入力するものであり、その出力は左(L)、右(R)チャネルのA/D変換器2a、2bを介してFIRフィルタ3a、3bに入力される。FIRフィルタ3a、3bは後述する信号処理を行うデジタルフィルタであり、夫々の出力はD/A変換器4a、4b、アンプ5a、5bを介して、ヘッドホンスピーカ6a、6bに与えられる。

【0011】さて図3に示す $h_1(t)$ は、ヘッドホンスピーカ6aと受聴者7の左耳の位置における頭部伝達関数である。正確には頭部伝達関数はヘッドホンスピーカにインパルスを入力した際の鼓膜の位置での応答である。以下、時間領域での説明を行うため伝達関数をインパルス応答と呼ぶ。但し、周波数領域で考えても同様の結果が得られる。

【0012】同様に $h_2(t)$ は、ヘッドホンスピーカ6bと受聴者7の右耳の位置におけるインパルス応答である。また信号発生手段8はインパルス信号等の信号S(t)を発生する手段であり、その出力は仮想スピーカ9に与えられる。仮想スピーカ9は受聴者7の左後方に配置された仮想のスピーカである。またFIRフィルタ3a、3bは入力信号を夫々インパルス応答 $h_{LL}(t)$ 、 $h_{RR}(t)$ で畳込み演算をするデジタルフィルタである。

【0013】ここで、音像を任意の方向に仮想的に定位させる場合について考える。左チャネルのヘッドホンスピーカ6aと右チャネルのヘッドホンスピーカ6bを用いて、左後方に設けられた仮想スピーカ9に相当する音を仮想的に生成する音場信号再生の原理について述

べる。まず左後方の音源に相当する音声信号 $S(t)$ が A/D 変換器 2 a、2 b に入力され、デジタル信号に変換される。ここで変換されたデジタル音声信号は FIR フィルタ 3 a、3 b に入力されるものとする。

【0014】これに対して信号発生手段 8 の信号 $S(t)$ を仮想スピーカ 9 から出力した場合、受聴者 7 の耳に到達する音は、次の式ようになる。すなわち、左耳における音圧 $L(t)$ は次の (1) 式で表される。

$$L(t) = S(t) * h_3(t) \cdots (1)$$

右耳における音圧 $R(t)$ は (2) 式で表される。

$$R(t) = S(t) * h_4(t) \cdots (2)$$

但し、* は畳込み演算を表している。実際は、スピーカ自身の伝達関数などが掛け合わされることとなるが、これは無視することとする。またスピーカ等の伝達関数が *

$$L(n) = S(n) * h_3(n)$$

$$= \sum_{k=0}^{N-1} S(k) \cdot h_3(n-k) \cdots (8)$$

【数 2】

$$R(n) = S(n) * h_4(n)$$

$$= \sum_{k=0}^{N-1} S(k) \cdot h_4(n-k) \cdots (9)$$

ここで、自然数 n は厳密には nT と表記すべきであり、 T はサンプリング周期を表すが、一般的に T を省略して (8)、(9) 式のように表記する。また、 N はインパルス応答 $h_3(n)$ 、 $h_4(n)$ の長さである。

【0017】また、同様に信号 $S(t)$ が信号入力手段 1 を介してヘッドホンスピーカ 6 a、6 b から音として出力されて、受聴者 7 に到達する音について、次の式が成立する。すなわち左耳の音圧は、次の (10) 式となる。

$$L'(n) = S(n) * h_{LL}(n) * h_1(n) \cdots (10)$$

右耳の音圧は、次の (11) 式となる。

$$R'(n) = S(n) * h_{RR}(n) * h_2(n) \cdots (11)$$

【0018】頭部伝達関数が互いに等しければ、一般に音が同一方向から聞こえるということを前提にする (この前提は一般的に正しい) と、次の (12) ~ (15) 式が成立する。

$$L(n) = L'(n) \cdots (12)$$

$$h_3(n) = h_{LL}(n) * h_1(n) \cdots (13)$$

$$R(n) = R'(n) \cdots (14)$$

$$h_4(n) = h_{RR}(n) * h_2(n) \cdots (15)$$

従って (13)、(15) 式が成立するように、インパルス応答 $h_{LL}(n)$ 、 $h_{RR}(n)$ を決定すれば良い。

【0019】例えば、インパルス応答 $h_1(n) \sim h_4(n)$ 、

* $h_3(t)$ 、 $h_4(t)$ に含まれていると考えても良い。

【0015】また、(1) (2) 式の音圧、インパルス応答、及び信号 $S(t)$ を時間軸上で離散的なデジタル信号として考え、夫々次の (3) ~ (7) 式のように変換する。

$$L(t) \rightarrow L(n) \cdots (3)$$

$$R(t) \rightarrow R(n) \cdots (4)$$

$$h_3(t) \rightarrow h_3(n) \cdots (5)$$

$$h_4(t) \rightarrow h_4(n) \cdots (6)$$

$$S(t) \rightarrow S(n) \cdots (7)$$

【0016】この場合、前述した (1) 式、(2) 式は次の (8)、(9) 式のように書き変えることができる。

【数 1】

$h_{LL}(n)$ 、 $h_{RR}(n)$ を周波数領域の表現で書き直すと、次の (16) ~ (21) 式ようになる。

$$H_1(n) = \text{FFT}(h_1(n)) \cdots (16)$$

$$H_2(n) = \text{FFT}(h_2(n)) \cdots (17)$$

$$H_3(n) = \text{FFT}(h_3(n)) \cdots (18)$$

$$H_4(n) = \text{FFT}(h_4(n)) \cdots (19)$$

$$H_{LL}(n) = \text{FFT}(h_{LL}(n)) \cdots (20)$$

$$H_{RR}(n) = \text{FFT}(h_{RR}(n)) \cdots (21)$$

但し、 $\text{FFT}()$ はフーリエ変換 (FFT) された関数を表す。

【0020】次に、(13)、(15) 式を周波数領域の表現で書き直すと、次に示す (22)、(23) 式のように畳込み演算が乗算に変わり、後はそれぞれのインパルス応答をフーリエ変換した伝達関数になる。

$$H_3(n) = H_{LL}(n) \cdot H_1(n) \cdots (22)$$

$$H_4(n) = H_{RR}(n) \cdot H_2(n) \cdots (23)$$

(22)、(23) 式において、FIR フィルタの伝達関数 $H_{LL}(n)$ 、 $H_{RR}(n)$ の値以外は測定により得られることから、次に示す (24)、(25) 式のように FIR フィルタの伝達関数 $H_{LL}(n)$ 、 $H_{RR}(n)$ を求めることができる。

【数 3】

$$HLL(n) = \frac{H3(n)}{H1(n)}$$

$$\dots\dots\dots (24)$$

【数 4】

$$HRR(n) = \frac{H4(n)}{H2(n)}$$

$$\dots\dots\dots (25)$$

【0021】このようにして決定された伝達関数HLL(n)、HRR(n)を逆フーリエ変換(IFFT)したhLL(n)、hRR(n)を用い、信号S(n)を図3のFIRフィルタ3a、3bで演算を行う。このため、ヘッドホンスピーカ6aに入力される信号はインパルス応答hLL(n)で畳み込まれ、ヘッドホンスピーカ6bに入力される信号はインパルス応答hRR(n)で畳み込まれる。こうして音声をヘッドホンスピーカ6a、6bから出力すると、受聴者7は実際に左後方の仮想スピーカ9を鳴らさなくても、その方向から音が鳴っていると感じることができる。このような方法により、音像を任意の方向に仮想的に定位させることができる。

【0022】次に、本発明の第1実施例における音場再生装置について図1、図2を用いて説明する。図1は第1実施例による音場再生装置の全体構成を示すブロック図であり、図2は信号処理手段3の内部構成を示すブロック図である。図1において信号入力手段1はアナログ(モノラル)の音声信号S(t)を入力し、A/D変換器2に出力する回路である。A/D変換器2はアナログの音声信号を周波数fsでサンプリングし、デジタルの音声信号に変換する回路である。なおここでのサンプリング周波数fsは44.1kHzに設定されることが多い。

【0023】A/D変換器2の出力は信号処理手段3に与えられる。信号処理手段3の制御信号入力端はパラメータ制御手段11を介して入力手段10に接続されている。入力手段10は、音場再生装置の使用である受聴者7の頭部形状や大きさ、受聴者7と音像の位置、音場の広がり等を入力する手段である。またパラメータ制御手段11は入力手段10から設定された条件に応じて信号処理手段3におけるパラメータの値(畳み込み係数)を制御するものである。パラメータ制御手段11は受聴者から見て全ての方向及び位置に音像を定位させるための畳み込み係数をパラメータとしてあらかじめ記憶しており、入力手段10の信号によりこれを選択して信号処理手段3に出力する。

【0024】図2に示すように、信号処理手段3は左右チャンネルの信号処理手段3c、3dから成り、左チャンネルの信号処理手段3cは音像定位手段12a、オーバーサンプル手段14a、フィルタ手段15a、ダウンサンプル手段16aとで構成される。また右チャンネルの信号処理手段3dは音像定位手段12b、オーバーサンプル手段14b、フィルタ手段15b、ダウンサンプル手段16bにより構成される。図1のパラメータ制御手段11の出力する各タップ係数(畳み込み係数)は音像

定位手段12a、12bに与えられる。

【0025】音像定位手段12a、12bは図3のFIRフィルタ3a、3bに相当するデジタルフィルタにより構成される。図2の音像定位手段12a、12bは図1のA/D変換器2から周波数fsでサンプリングされた音声信号を入力し、音像の位置に対応した畳み込み係数で演算を行う。このとき音像定位手段12a、12bはパラメータ制御手段11から与えられるパラメータをFIRフィルタの一部を構成する遅延回路及びその出力が入力される各乗算器に対応させて、時系列の信号に変換する。ここで時系列に並べられたパラメータをオーバーサンプル手段14a、14bに与え、パラメータをn×fsの周波数でオーバーサンプリングする。そしてオーバーサンプリングされたパラメータはフィルタ手段15a、15bに与えられる。

【0026】フィルタ手段15a、15bはfs/2以下の遮断周波数を有するローパスフィルタである。ダウンサンプル手段16a、16bは、フィルタ手段15a、15bの信号を夫々入力し、m×fs(m>1)の周波数でサンプリングする回路である。ダウンサンプル手段16a、16bの出力は再び時系列の信号から夫々のパラメータに変換されて音像定位手段12a、12bに入力される。そして音像定位手段12a、12bは、パラメータ制御手段11を介して入力手段10から入力された受聴者の標準の頭部形状に対して、個人差のデータが与えられると、補正された畳み込み係数で演算を行い、その出力を夫々図1のD/A変換器4a、4bに出力する。

【0027】このように構成された音場再生装置の動作について説明する。まず図1の信号入力手段1には音声(オーディオ)信号が入力される。この音声信号はA/D変換器2でデジタル信号に変換され、信号処理手段3に入力される。入力手段10で受聴者の頭部形状や大きさ、受聴者と音像の位置や距離、音場の広さ等を入力する。パラメータ制御手段11は入力手段10で入力された内容に応じた特性が得られるようパラメータを選択する。すなわちパラメータ制御手段11は、あらかじめ記憶している畳み込み係数を選択し、信号処理手段3に出力する。

【0028】ある音場の受聴者7が、図4に示すようにある方向の音を聴く場合、両耳に入力されるインパルス応答は例えば図5(a)のようになり、周波数特性は図5(b)のようになる。そして受聴者の頭部形状や大きさ、受聴者と音像の位置関係に応じて、左右両耳の音量

バランス、到達時間、位相が変わることになる。まず、受聴者 7 の頭部形状や受聴者と音像との位置関係を図 1 の入力手段 10 によって設定する。パラメータ制御手段 11 は入力された音像位置に応じてあらかじめ記憶された畳み込み係数を選択し、図 2 の音像定位手段 12 a、12 b に与える。このとき出力される係数には頭部個人差は考慮されていない一般的な値である。音像定位手段 12 a、12 b は前述した (24) 式、(25) 式で示す伝達関数で演算を行う。このときオーバーサンプル手段 14 a、14 b は、標準的な頭部形状と受聴者の頭部形状との差をもとに、係数の時間軸をずらすためにオーバーサンプルをする。

【0029】図 6 はオーバーサンプルからダウンサンプルまでの概念を示す一般的な説明図である。図 6 (a) は入力された制御信号 (パラメータ) を時系列の信号に変換してサンプリング周波数 f_s でサンプリングした場合の波形である。(a) の信号をオーバーサンプリングするため、図 6 (b) に示すように (a) で示す信号のサンプリング周期を更に n 等分する。こうして得られた $1/(n \times f_s)$ の周期で信号間を補間すると、図 6 (c) のように細かい間隔の信号となる。

【0030】この信号を図 2 のフィルタ手段 15 a、15 b に与え、補間された信号にカットオフ周波数 $f_s/2$ のローパスフィルタに入力すると、サンプリングホールドされて図 6 (d) のような滑らかな曲線の信号が出力される。この信号をダウンサンプル手段 16 a、16 b に与えると、 $m \times f_s$ ($n > m > 1$) の周波数でサンプリングが行われ、図 6 (e) のような信号が生成される。図 6 (e) に示す信号は (a) に示す信号より、同一時間間隔でデータの個数が $14/11$ 倍に増加している。こうして元のサンプリング波形より細かい時系列の制御信号が得られる。そしてこの時系列の信号を例えばシリアルイン・パラレルアウトのシフトレジスタに与え、各信号を同時に取り出すと、所望の値に補正された制御信号が得られる。

【0031】次に、この手法を畳み込み係数に適用し、より正確な音像定位を実現する場合について説明する。図 7 は畳み込み係数の個人差に対するデータのずらし方の概念を示している。図 7 (a) は入力音声信号に対する多くの受聴者から統計的に計算された代表的な畳み込み係数を示している。従ってこの係数を用いて音像定位手段 12 a、12 b で畳み込みを行えば、かなりの受聴者が自然な音像定位感を知覚できるはずである。しかし実際多くの場合、受聴者によっては頭部形状がばらつき、大きかったり小さかったりする。このため標準の係数では両耳間隔に微妙な差が生じ、各受聴者に対して補正が必要となる。

【0032】この補正方法として、係数を時間的にずらして左右両耳に到達する音の位相差を変える方法をとる。図 7 (a) の係数を仮に 1 タップずらすと (b) の

ようになる。例えばサンプリング周波数 f_s を 44.1kHz とし、この周波数でサンプリングすると、係数を 1 タップずらしたとしても両耳間隔差は 7.7mm が限界となる。従って 7.7mm 以下の両耳間隔を有する個人差に対応できなくなる。しかし、 n 倍にオーバーサンプルすると、 $7.7/n$ (mm) 程度の個人差にも対応できることとなる。

【0033】図 7 (c) は (a) に示す係数をオーバーサンプルしたものである。この状態で 1 タップずらしていくと図 7 (d) ~ (f) のように、同じ係数でもより細かく位相をずらすことができる。特にインナホンのように、個人の微妙な両耳間隔差が音像定位感に大きく影響する場合には有効な方法となる。このように任意の受聴者に対して音像を希望する方向に定位させることができる。

【0034】つぎに本発明の第 2 実施例 (請求項 3) における音場再生装置について図 8 を用いて説明する。図 8 は本実施例の音場再生装置の一部の構成を示すブロック図であり、図 2 に示す信号処理手段 3 c、3 d の後段に付加される回路だけを示している。その他のブロックは第 1 実施例と同一であり、全体構成の説明は省略する。図 8 に示すように信号処理手段 3 c、3 d の出力信号に対し、残響音作成手段 17 a、17 b が設けられている。

【0035】残響音作成手段 17 a、17 b は音場の広さに応じた残響音を作成し、付加する手段であり、その出力は図 1 の D/A 変換器 4 a、4 b に夫々与えられる。残響音作成手段 17 a、17 b は、例えば遅延時間の異なる複数のフィードバックエコー回路を直列接続することにより構成される。残響音として例えば 10m 前後までの空間の広がりを出す場合には、残響音の長さは例えば 0.25 ~ 0.35 秒となる。また 10m 以上 20m 前後の距離感に空間の広がりを出す場合には、例えば残響音の長さは 0.7 ~ 0.9 秒に設定する。又、広いコンサートホールのような音場再生ならば残響時間を比較的長く、かつ低音域の残響時間を高音域のそれより長くするように付加する。

【0036】図 9 は第 3 実施例 (請求項 4) における音場再生装置のパラメータ制御の概念図である。本実施例の音場再生装置の構成も図 1 と同一である。本実施例では入力手段 10 によって音源の位置や音場の広さを入力すると共に、受聴者 P と音像 S との距離関係を示す側方反射角度 θ を入力する。例えば図 9 の受聴者 P 2 で示すように音像 S からの距離が遠くなると、 θ の値が小さくなる。さらに受聴者 P 1 で示すように距離が近くなると、 θ の値は大きくなる。この側方反射角 θ の大きさによって信号処理手段 3 における遅延や畳み込み係数が変わるものとする。

【0037】図 10 は本発明の第 4 実施例 (請求項 5) における音場再生装置のブロック図であり、第 1 実施例と同一部分は同一の符号をつけ、それらの説明は省略す

る。本図に示すように図1の入力手段10に代わってパラメータ受信手段18が設けられている。パラメータ受信手段で18は外部から位置や距離や広がり感を制御するための制御信号を受信する手段である。

【0038】パラメータ受信手段18では、外部から位置や距離や広がり感の条件を含んだ制御信号を受信すると、その制御信号をパラメータ制御手段11に与える。そしてパラメータ制御手段11は信号処理手段3で設定すべきパラメータ（各係数）を出力する。以下の動作は前述した第1実施例と同様である。

【0039】以上のように各実施例では、任意の受聴者に対してあらかじめ設定された音像定位のための畳み込み係数をオーバーサンプルすることによって、細かく位相調整して対応することができる。そして、外部からの制御信号により音の定位感、距離感や広がり感を制御することができ、あらかじめプログラミングされた信号を用いて、繰り返しの制御を行ったり、映像などと組み合わせて、映像の場面に応じた音の定位感、距離感、広がり感の制御が可能となる。

【0040】なお、本実施例では入力信号がモノラルの音声信号の場合について述べたが、ステレオの場合でも、信号毎に独立の信号処理手段を設けることで対応でき、またヘッドホンやインナホンの代わりに2チャンネルスピーカにおいても有効に適用される。

【0041】

【発明の効果】以上のように本発明によれば、入力手段に入力された受聴者と音像の位置、受聴者の頭部の形状と大きさ、距離や音場の広さに応じてパラメータを制御し、音像を定位させるための畳み込み係数をオーバーサンプルすることによって、任意の受聴者に対して個人の微妙な両耳間隔差に対応でき自然な音像定位と音場再生が実現できる。

【図面の簡単な説明】

【図1】本発明の第1実施例における音場再生装置のブロック図である。

【図2】第1実施例の音場再生装置に用いられる信号処理手段の構成を示すブロック図である。

*

* 【図3】本発明の音場再生装置における音像定位の原理を示す説明図である。

【図4】受聴者と音像との関係を示す説明図である。

【図5】両耳の周波数特性とインパルス応答を示すグラフである。

【図6】オーバーサンプルからダウンサンプルまでの動作原理を示す概念図である。

【図7】受聴者の個人差に対する畳み込み係数のずらし方の概念図である。

10 【図8】本発明の第2実施例における音場再生装置の一部の構成を示すブロック図である。

【図9】本発明の第3実施例の音場再生装置におけるパラメータ制御の概念図である。

【図10】本発明の第4実施例における音場再生装置の構成を示すブロック図である。

【図11】従来の音像定位装置の構成例を示すブロック図である。

【符号の説明】

1 信号入力手段

2, 2 a, 2 b, A/D変換器

3, 3 c, 3 d 信号処理手段

3 a, 3 b FIRフィルタ

4 a, 4 b D/A変換器

5 a, 5 b アンプ

6 a, 6 b スピーカ

7 受聴者

8 信号発生手段

9 仮想スピーカ

10 入力手段

11 パラメータ制御手段

12 a, 12 b 音像定位手段

14 a, 14 b オーバーサンプル手段

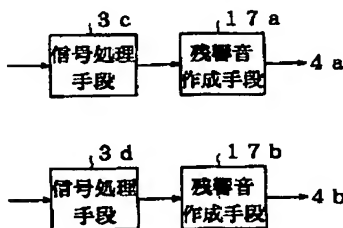
15 a, 15 b フィルタ手段

16 a, 16 b ダウンサンプル手段

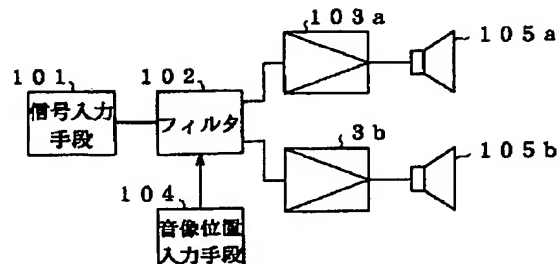
17 a, 17 b 残響音作成手段

18 パラメータ受信手段

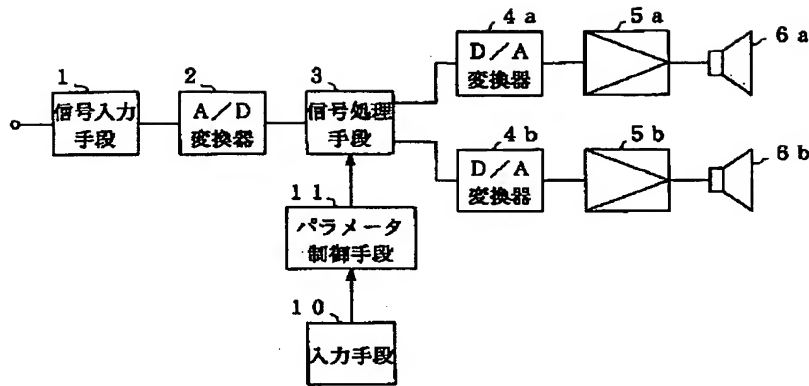
【図8】



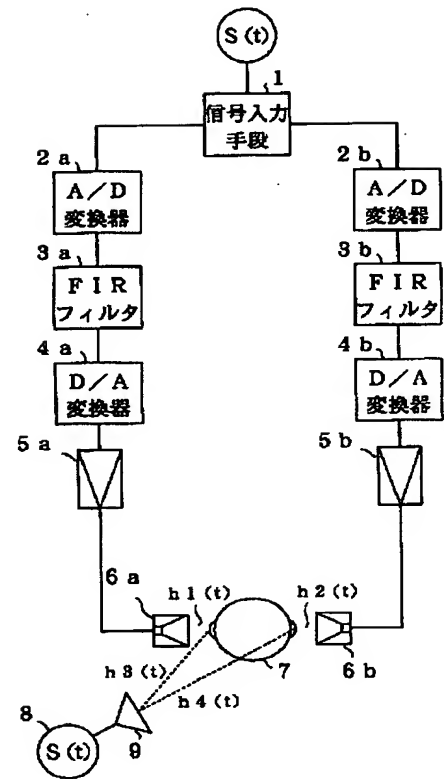
【図11】



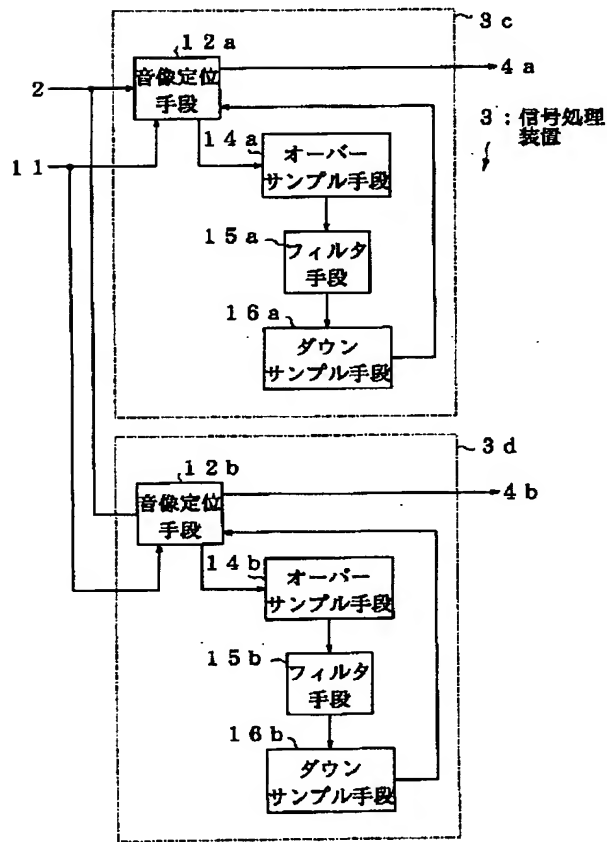
【図 1】



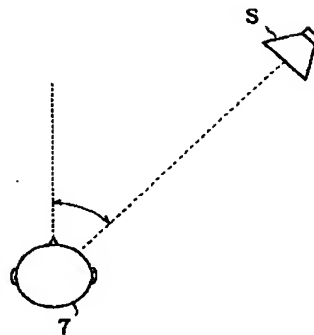
【図 3】



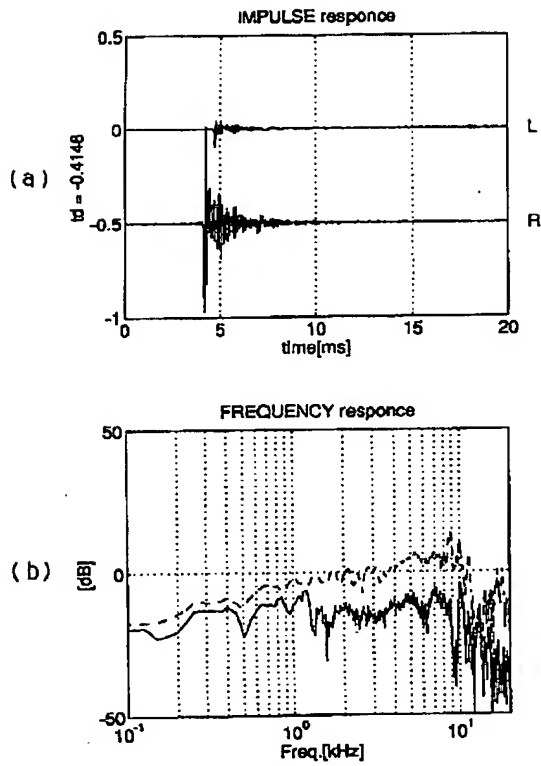
【図 2】



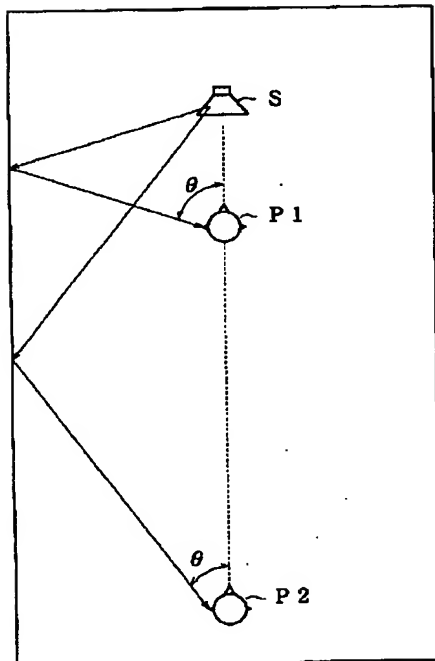
【図 4】



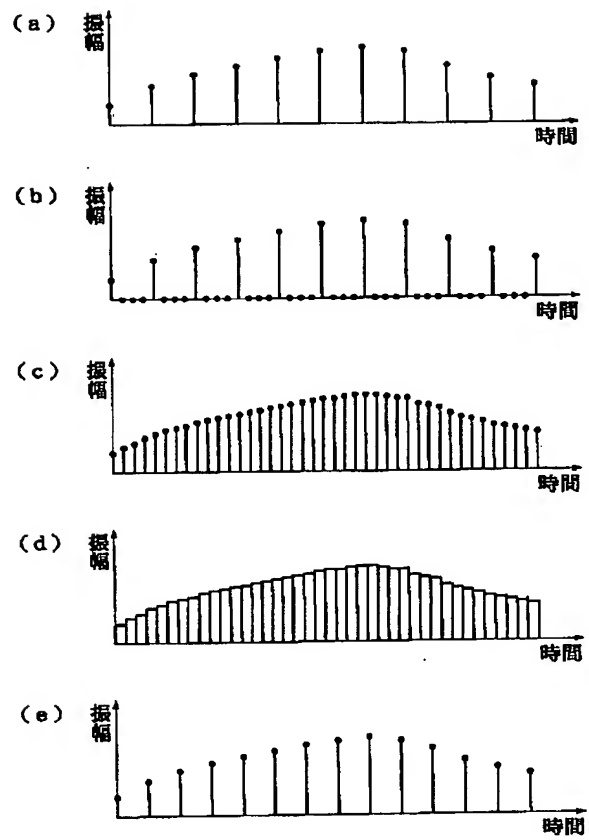
【図 5】



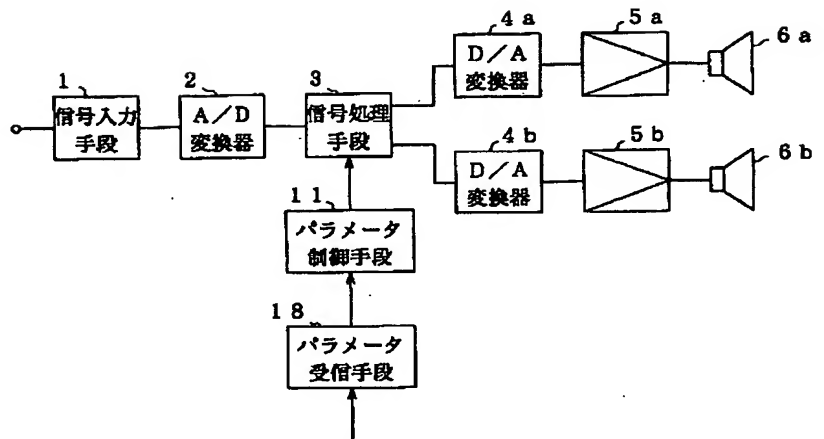
【図 9】



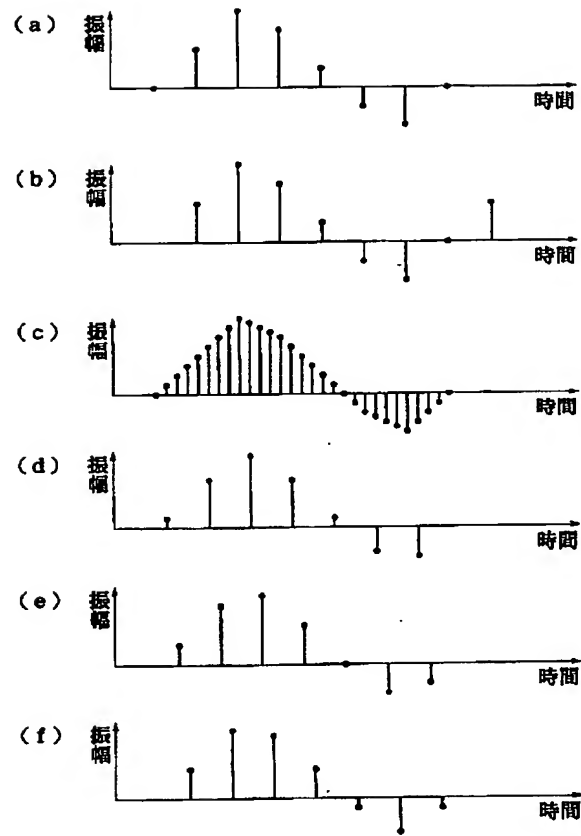
【図 6】



【図 10】



【図 7】



フロントページの続き

(72)発明者 川村 明久
 大阪府門真市大字門真1006番地 松下電器
 産業株式会社内

(72)発明者 松本 正治
 大阪府門真市大字門真1006番地 松下電器
 産業株式会社内

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-191225

(43)Date of publication of application : 23.07.1996

(51)Int.Cl. H03G 5/24

G10K 15/00

H04S 1/00

H04S 7/00

(21)Application number : 07-018576 (71)Applicant : MATSUSHITA

ELECTRIC IND CO LTD

(22)Date of filing : 09.01.1995 (72)Inventor : OGAWA MASAKO

TAMURA TADASHI

KATAYAMA TAKASHI

KAWAMURA AKIHISA

MATSUMOTO MASA HARU

(54) SOUND FIELD REPRODUCING DEVICE

(57)Abstract:

PURPOSE: To attain natural senses of sound image positioning, distance and expansion in all directions by the use of a headphone or an innerphone.

CONSTITUTION: A sound signal from a signal input means 1 is applied to an A/D converter 2, which converts the sound signal into a digital signal at a sampling frequency f_s . At the time of inputting the digital signal, a signal processing means 3 executes operation for sound image positioning by the use of an FIR filter. In this case, a convolution coefficient is over-sampled and a tap position for multiplying the convolution coefficient is finely controlled in

accordance with a distance difference between both ears of a listener. Thus the convolution coefficient is controlled in accordance with the shape and size of the head part of the listener and positional relation between the listener and a sound image. Thus the operated sound signal is applied to D/A converters 4a, 4b and D/A converted signals are outputted from speakers 6a, 6b of a headphone through amplifiers 5a, 5b.

LEGAL STATUS [Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

*** NOTICES ***

**JPO and NCIP are not responsible for any
damages caused by the use of this translation.**

**1. This document has been translated by computer. So the translation may not
reflect the original precisely.**

2. ** shows the word which can not be translated.**

3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

**[Claim 1] The digital signal changed by the A/D converter which changes an
input sound signal into a digital signal, and said A/D converter is inputted. A
signal-processing means to perform signal processing so that it may become an
acoustic signal according to the location and distance of a listening person's
head configuration, magnitude, a listening person, and an image, A listening
person's head configuration, magnitude, and an input means to input the
location of a listening person and an image, A parameter-control means to set
up a parameter so that the acoustic feature according to the content inputted by**

said input means may be acquired, and to give said parameter to said signal-processing means, The D/A converter which changes into an analog signal the signal outputted from said signal-processing means, It is a sound field regenerative apparatus possessing the amplifier which amplifies the output of said D/A converter, and the sound transducer of the right and left which carry out the alteration in percussion sound of the output of said amplifier. Said signal-processing means An image normal means to make an image orientate by the reefing operation according to the physical relationship of the listening person and image which were inputted by said input means, The reefing multiplier of criteria is inputted as a parameter from said parameter-control means through said image normal means. The exaggerated sample means which changes these parameters into the discretization signal of time series, and carries out the exaggerated sample of said parameter with the value more than the sampling frequency of said A/D converter, A filter means to pass the component below predetermined frequency characteristics to the parameter processed with said exaggerated sample means, The down sample of the parameter outputted from said filter means is carried out with the value below the sampling frequency of said exaggerated sample means. The sound field regenerative apparatus which carries out the parallel conversion of the signal of the acquired time series to each reefing multiplier, and is characterized by being

what has a down sample means which is collapsed and is given to said image normal means as a multiplier by which the value was amended.

[Claim 2] Said filter means is a sound field regenerative apparatus according to claim 1 characterized by being what passes $1/2$ or less frequency component of the sampling frequency of said A/D converter.

[Claim 3] Said input means is a sound field regenerative apparatus according to claim 1 characterized by establishing a reverberation sound creation means to apply the reverberation sound corresponding to the size of the sound field where the breadth of sound field is also inputted simultaneously and is obtained by the output section of said signal-processing means from said input means while inputting the location of a listening person's head configuration, magnitude and a listening person, and an image.

[Claim 4] Said parameter-control means is a sound field regenerative apparatus according to claim 1 characterized by to be what changes said include angle into the listening location or image location in sound space, and gives a parameter to said signal-processing means, when the sound source and the direction of a reflected sound which were seen from the listening person, and an include angle are inputted in the reflected sound by which reflects with a wall from the sound source of sound space, and incidence is carried out to a listening person.

[Claim 5] Said input means is a sound field regenerative apparatus according to

claim 1 characterized by being a thing including the parameter receiving means which receives the control signal about sound field from the exterior, and is given to said parameter-control means.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] In AV (audio-visual) device, this invention relates to the sound field regenerative apparatus in consideration of both individual lugs spacing and a head configuration, in order to perform sound reproduction which has a natural feeling of normal, and presence to the listening person of arbitration.

[0002]

[Description of the Prior Art] In recent years, in the multimedia field consisting mainly of an image and sound, development of the hardware corresponding to the three-dimensional sound reproduction of a three dimension is desired with the formation of a three-dimension computer graphic of an image. Moreover, an interactive device like the home video game machine for TV, a personal

computer, etc. is spreading. Not only a two-channel stereo loudspeaker but development which controls interactively the feeling of normal of a sound, a sense of distance, and a feeling of breadth in the sound field playback using a year loudspeaker, headphone, and an inner phone is performed that it should correspond to this.

[0003] Drawing 11 is the block diagram showing an example of the conventional sound field regenerative apparatus which used the loudspeaker. In this drawing, the signal input means 101 is an input means to input an audio signal, and the output is given to Amplifier 103a and 103b through a filter 102. A filter 102 controls the frequency characteristics of 2 channel signals, sound volume, and time delay, and is controlled by the image location input means 104. The image location input means 104 is a means to input an image location, in order to control the feeling of normal of a listening person's sound. The sound signal amplified with Amplifier 103a and 103b is given to Loudspeakers 105a and 105b.

[0004] Thus, actuation of the constituted conventional sound field regenerative apparatus is explained. The sound signal inputted by the signal input means 101 controls a filter 102 to reproduce voice from Loudspeakers 105a and 105b by the frequency characteristics according to the location of **** inputted with the image location input means 104, sound volume, and time delay. a right ear -- the direction of a left ear -- time delay -- large -- the crown -- if sound pressure level

becomes low with the frequency characteristics of a region, it will be sensed general that an image is in the method of the right. A filter 102 performs such control and controls a listening person's feeling of image normal. The signal controlled by the filter 102 is amplified with Amplifier 103a and 103b, and is reproduced by Loudspeakers 105a and 105b. By performing such signal processing, the feeling of normal of a sound is controllable.

[0005]

[Problem(s) to be Solved by the Invention] However, it was difficult to control change of the delicate property by difference of the configuration of a listening person's head and a handle part by the above configurations, even if it hears a playback sound. In the case of the inner phone inserted especially in headphone and external auditory meatus, the feeling of the outside of the head of the front in an image, a feeling of normal, and a sense of distance were unrealizable.

[0006] This invention is made in view of such a conventional trouble, and aims at offering a sound field regenerative apparatus with which the natural feeling of image normal of all directions and a sense of distance are acquired to the listening person of arbitration.

[0007]

[Means for Solving the Problem] The digital signal from which this invention was changed by the A/D converter which changes an input sound signal into a digital

signal, and the A/D converter is inputted. A signal-processing means to perform signal processing so that it may become an acoustic signal according to the location and distance of a listening person's head configuration, magnitude, a listening person, and an image, A listening person's head configuration, magnitude, and an input means to input the location of a listening person and an image, A parameter-control means to set up a parameter so that the acoustic feature according to the content inputted by the input means may be acquired, and to give a parameter to a signal-processing means, The D/A converter which changes into an analog signal the signal outputted from a signal-processing means, It is a sound field regenerative apparatus possessing the amplifier which amplifies the output of a D/A converter, and the sound transducer of the right and left which carry out the alteration in percussion sound of the output of amplifier. A signal-processing means An image normal means to make an image orientate by the reefing operation according to the physical relationship of the listening person and image which were inputted by the input means, The reefing multiplier of criteria is inputted as a parameter from a parameter-control means through an image normal means. The exaggerated sample means which changes these parameters into the discretization signal of time series, and carries out the exaggerated sample of the parameter with the value more than the sampling frequency of an A/D converter, A filter means to pass the

component below predetermined frequency characteristics to the parameter processed with the exaggerated sample means, The down sample of the parameter outputted from a filter means is carried out with the value below the sampling frequency of an exaggerated sample means. The parallel conversion of the signal of the acquired time series is carried out to each reeving multiplier, and it is characterized by having a down sample means which is collapsed and is given to an image normal means as a multiplier by which the value was amended.

[0008]

[Function] According to this invention which has such a description, input setting out of a listening person's head configuration, magnitude, and the location and distance of a listening person and an image is carried out with an input means. If a sound signal is inputted through a signal input means, it will be changed into a digital signal by the A/D converter, and will be inputted into a signal-processing means. A parameter is chosen as a parameter-control means so that the acoustic feature according to the content set up by the input means may be acquired. And PAMETA corresponding to institution conditions is collapsed and it gives as a multiplier a signal-processing means. The signal calculated by the digital disposal circuit is changed into an analog signal by the D/A converter, and a sound is outputted from a sound transducer through amplifier.

[0009] After carrying out image normal according to the physical relationship of a listening person's head configuration, magnitude, a listening person, and an image inputted [especially] by the input means with the signal-processing means, carrying out the exaggerated sample of the output and filtering with predetermined frequency characteristics, he carries out a down sample with the value near the original sampling frequency, and is trying to output an acoustic signal. In order to make both the delicate lugs spacing difference of the listening person of arbitration reflect in an output signal at this time, an exaggerated sample is carried out with a required sampling frequency. Moreover, if the signal by which over sampling technique was carried out is given to a filter means, the thing of $1/2$ or less component of the sampling frequency of an A/D converter will pass. If it carries out like this, the distortion at the time of a down sample being carried out with a down sampling means will decrease. Thus, if it processes, it will correspond to both the delicate lugs spacing difference to the listening person of arbitration, and natural image normal and sound field playback will be obtained to all directions.

[0010]

[Example] It explains referring to a drawing about the sound field regenerative apparatus in the example of this invention. How to make a sound orientate behind [left] the listening person 7 from the headphone loudspeakers 6a and 6b

is first explained using principle drawing of the image normal of drawing 3 .

Setting to this Fig., the signal input means 1 is sound signal $S(t)$. It inputs and the output is inputted into the FIR filters 3a and 3b through A/D-converter 2a of (Left L) (right R) tea YANERU, and 2b. The FIR filters 3a and 3b are digital filters which perform signal processing mentioned later, and each output is given to the headphone loudspeakers 6a and 6b through D/A converters 4a and 4b and Amplifier 5a and 5b.

[0011] Now, $h_1(t)$ shown in drawing 3 is a head transfer function in the location of the left ear of headphone loudspeaker 6a and the listening person 7. It is a response in the location of the eardrum at the time of a head transfer function inputting an impulse into accuracy at a headphone loudspeaker. Hereafter, a transfer function is called an impulse response in order to give explanation in a time domain. However, the same result is obtained even if it thinks in a frequency domain.

[0012] $h_2(t)$ is an impulse response in the location of the right ear of headphone loudspeaker 6b and the listening person 7 similarly. Moreover, the signal generation means 8 is signals [, such as an impulse signal,] $S(t)$. It is a means to generate and the output is given to the virtual loudspeaker 9. The virtual loudspeaker 9 is a loudspeaker of the imagination arranged behind [left] the listening person 7. Moreover, the FIR filters 3a and 3b are impulse responses

$h_{LL}(t)$ about an input signal, respectively. $h_{RR}(t)$ It is the digital filter which calculates by collapsing.

[0013] Here, the case where an image is made to orientate in the direction of arbitration virtually is considered. The principle of the sound field signal regeneration which generates virtually the sound equivalent to the virtual loudspeaker 9 prepared in left back is described using headphone loudspeaker 6a of a left channel, and headphone loudspeaker 6b of a right channel. Sound signal $S(t)$ which is equivalent to the sound source of left back first is inputted into A/D-converter 2a and 2b, and is changed into a digital signal. The digitized voice signal changed here shall be inputted into the FIR filters 3a and 3b.

[0014] On the other hand, signal [of the signal generation means 8] $S(t)$ When it outputs from the virtual loudspeaker 9, the sound which reaches the listening person's 7 lug becomes like the following formula. That is, sound pressure $L(t)$ in a left ear is expressed with the following (1) type.

$$L(t) = S(t) * h_3(t) \dots (1)$$

Sound pressure R in a right ear (t) It is expressed with (2) types.

$$R(t) = S(t) * h_4(t) \dots (2)$$

However, $*$ collapses and expresses the operation. In practice, although the own transfer function of a loudspeaker etc. will be multiplied, this decides to ignore.

Moreover, you may think that transfer functions, such as a loudspeaker, are

contained in $h_3(t)$ and $h_4(t)$.

[0015] Moreover, the sound pressure of (1) and (2) type, an impulse response, and signal $S(t)$ On a time-axis, it thinks as a discrete digital signal and changes like the following (3) - (7) type, respectively.

$$L(t) \rightarrow L(n) \dots (3)$$

$$R(t) \rightarrow R(n) \dots (4)$$

$$h_3(t) \rightarrow h_3(n) \dots (5)$$

$$h_4(t) \rightarrow h_4(n) \dots (6)$$

$$S(t) \rightarrow S(n) \dots (7)$$

[0016] In this case, (1) type and (2) types which were mentioned above can be written like the following (8) and (9) types, and can be changed.

[Equation 1]

$$\begin{aligned} L(n) &= S(n) * h_3(n) \\ &= \sum_{k=0}^{N-1} S(k) \cdot h_3(n-k) \dots\dots\dots (8) \end{aligned}$$

[Equation 2]

$$\begin{aligned} R(n) &= S(n) * h_4(n) \\ &= \sum_{k=0}^{N-1} S(k) \cdot h_4(n-k) \dots\dots\dots (9) \end{aligned}$$

Here, although it should be strictly written as nT and T expresses a sampling period, generally the natural number n omits T and writes it like (8) and (9) types.

Moreover, N is an impulse response $h_3(n)$. It is the die length of $h_4(n)$.

[0017] Moreover, it is signal $S(t)$ similarly. It is outputted as a sound from the headphone loudspeakers 6a and 6b through the signal input means 1, and the following formula is materialized about the sound which reaches the listening person 7. That is, the sound pressure of a left ear serves as the following (10) types.

$$L'(n) = S(n) * h_{LL}(n) * h_1(n) \dots (10)$$

The sound pressure of a right ear serves as the following (11) types.

$$R'(n) = S(n) * h_{RR}(n) * h_2(n) \dots (11)$$

[0018] a head transfer function -- mutual -- etc. -- a premise [say / that a sound can generally be heard from the same if damp] -- carrying out (this premise being the right generally) -- the following (12) - (15) type is materialized.

$$L(n) = L'(n) \dots (12)$$

$$h_3(n) = h_{LL}(n) * h_1(n) \dots (13)$$

$$R(n) = R'(n) \dots (14)$$

$$h_4(n) = h_{RR}(n) * h_2(n) \dots (15)$$

Therefore, what is necessary is just to determine an impulse response $h_{LL}(n)$ and $h_{RR}(n)$ so that (13) and (15) types may be materialized.

[0019] for example, impulse response $h_1(n) - h_4(n)$ If $h_{LL}(n)$ and $h_{RR}(n)$ are rewritten with the expression of a frequency domain, it will become like the

following (16) - (21) type.

$$H1(n) = \text{FFT}(h1(n)) \dots (16)$$

$$H2(n) = \text{FFT}(h2(n)) \dots (17)$$

$$H3(n) = \text{FFT}(h3(n)) \dots (18)$$

$$H4(n) = \text{FFT}(h4(n)) \dots (19)$$

$$HLL(n) = \text{FFT}(hLL(n)) \dots (20)$$

$$HRR(n) = \text{FFT}(hRR(n)) \dots (21)$$

However, FFT () The function by which the Fourier transform (FFT) was carried out is expressed.

[0020] Next, if (13) and (15) types are rewritten with the expression of a frequency domain, it collapses like (22) and (23) types which are shown below, and an operation will change to multiplication and the rest will become the transfer function which carried out the Fourier transform of each impulse response.

$$H3(n) = HLL(n), H1(n) \dots (22)$$

$$H4(n) = HRR(n), H2(n) \dots (23)$$

In (22) and (23) types, except the transfer function HLL of an FIR filter (n), and the value of HRR (n), since it is obtained by measurement, the transfer function HLL of an FIR filter (n) and HRR (n) can be calculated like (24) and (25) types which are shown below.

[Equation 3]

$$H_{LL}(n) = \frac{H_3(n)}{H_1(n)} \quad \text{----- (24)}$$

[Equation 4]

$$H_{RR}(n) = \frac{H_4(n)}{H_2(n)} \quad \text{----- (25)}$$

[0021] Thus, $h_{LL}(n)$ which carried out the inverse Fourier transform (IFET) of the determined transfer function $H_{LL}(n)$ and the $H_{RR}(n)$, and $h_{RR}(n)$ are used, and it is signal $S(n)$. It calculates with the FIR filters 3a and 3b of drawing 3. For this reason, the signal inputted into headphone loudspeaker 6a is collapsed by the impulse response $h_{LL}(n)$, and the signal inputted into headphone loudspeaker 6b is collapsed by the impulse response $h_{RR}(n)$. In this way, if voice is outputted from the headphone loudspeakers 6a and 6b, even if the listening person 7 does not sound the virtual loudspeaker 9 of left back actually, he can sense that the sound is sounding from the direction. An image can be made to orientate in the direction of arbitration virtually by such approach.

[0022] Next, the sound field regenerative apparatus in the 1st example of this invention is explained using drawing 1 and drawing 2. Drawing 1 is the block diagram showing the whole sound field regenerative-apparatus configuration by the 1st example, and drawing 2 is the block diagram showing the internal

configuration of the signal-processing means 3. Setting to drawing 1 , the signal input means 1 is sound signal [of an analog (monophonic recording)] $S(t)$. It is the circuit which is inputted and is outputted to A/D converter 2. A/D converter 2 is a circuit which samples the sound signal of an analog on a frequency f_s , and is changed into a digital sound signal. In addition, the sampling frequency f_s here is 44.1kHz. It is set up in many cases.

[0023] The output of A/D converter 2 is given to the signal-processing means 3. The control signal input edge of the signal-processing means 3 is connected to the input means 10 through the parameter-control means 11. The input means 10 is a means to input the location of the head configuration of the listening person 7 who is a user of a sound field regenerative apparatus, magnitude, the listening person 7, and an image, the breadth of sound field, etc. Moreover, the parameter-control means 11 controls the value (reeling multiplier) of the parameter in the signal-processing means 3 according to the conditions set up from the input means 10. The parameter-control means 11 has memorized beforehand the convolution multiplier for seeing from a listening person and making all directions and locations orientate an image as a parameter, chooses this with the signal of the input means 10, and outputs it to the signal-processing means 3.

[0024] As shown in drawing 2 , the signal-processing means 3 consists of the

signal-processing means 3c and 3d of a right-and-left channel, and signal-processing means 3c of a left channel consists of image normal means 12a, exaggerated sample means 14a, filter means 15a, and down sample means 16a. Moreover, 3d of signal-processing means of a right channel is constituted by image normal means 12b, exaggerated sample means 14b, filter means 15b, and down sample means 16b. Each tap multiplier (reefing multiplier) which the parameter-control means 11 of drawing 1 outputs is given to the image normal means 12a and 12b.

[0025] The image normal means 12a and 12b are constituted by the digital filter equivalent to the FIR filters 3a and 3b of drawing 3 . The image normal means 12a and 12b of drawing 2 input the sound signal sampled on the frequency f_s from A/D converter 2 of drawing 1 , and it calculates by the convolution multiplier corresponding to the location of an image. At this time, the image normal means 12a and 12b make the parameter given from the parameter-control means 11 correspond to each multiplier into which the delay circuit which constitutes some FIR filters, and its output are inputted, and are changed into the signal of time series. The parameter put in order by time series here is given to the exaggerated sample means 14a and 14b, and over sampling technique of the parameter is carried out on the frequency of $nxfs$. And the parameter by which over sampling technique was carried out is given to the filter means 15a and 15b.

[0026] The filter means 15a and 15b are low pass filters which have $f_s / 2$ or less cut-off frequency. The down sample means 16a and 16b are circuits which input the signal of the filter means 15a and 15b, respectively, and are sampled on the frequency of mf_s ($m > 1$). The output of the down sample means 16a and 16b is again changed into each parameter from the signal of time series, and is inputted into the image normal means 12a and 12b. And if the data of individual difference are given to the head configuration of a listening person's criterion where it was inputted from the input means 10 through the parameter-control means 11, the image normal means 12a and 12b will calculate by the amended convolution multiplier, and will output the output to D/A converters 4a and 4b of drawing 1 , respectively.

[0027] Thus, actuation of the constituted sound field regenerative apparatus is explained. A voice (audio) signal is first inputted into the signal input means 1 of drawing 1 . This sound signal is changed into a digital signal with A/D converter 2, and is inputted into the signal-processing means 3. The size of a listening person's head configuration, magnitude, the location of a listening person and an image and distance, and sound field etc. is inputted with the input means 10. The parameter-control means 11 chooses a parameter so that the property according to the content inputted with the input means 10 may be acquired. Namely, the parameter-control means 11 chooses the convolution multiplier

memorized beforehand, and outputs it to the signal-processing means 3.

[0028] As shown in drawing 4 , when the listening person 7 of a certain sound field hears the sound of a certain direction, the impulse response inputted into both lugs becomes like drawing 5 (a), and frequency characteristics become like drawing 5 (b). And according to the physical relationship of a listening person's head configuration, magnitude, a listening person, and an image, the sound-volume balance of right-and-left both lugs, time of concentration, and a phase will change. First, the physical relationship of the listening person's 7 head configuration, or a listening person and an image is set up with the input means 10 of drawing 1 . The parameter-control means 11 chooses the convolution multiplier beforehand memorized according to the inputted image location, and gives it to the image normal means 12a and 12b of drawing 2 . It is the general value as which head individual difference is not considered by the multiplier outputted at this time. The image normal means 12a and 12b calculate with the transfer function shown by (24) types mentioned above and (25) formulas. At this time, based on the difference of a standard head configuration and a listening person's head configuration, the exaggerated sample means 14a and 14b carry out an exaggerated sample, in order to shift the time-axis of a multiplier.

[0029] Drawing 6 is the common explanatory view showing the concept from an

exaggerated sample to a down sample. Drawing 6 (a) is a wave at the time of changing the inputted control signal (parameter) into the signal of time series, and sampling with a sampling frequency f_s . In order to carry out the exaggerated sampling of the signal of (a), n division into equal parts of the sampling period of the signal shown by (a) as shown in drawing 6 (b) is done further. In this way, if between signals is interpolated with the period of obtained $1/(n \times f_s)$, it will become the signal of fine spacing like drawing 6 (c).

[0030] If this signal is given to the filter means 15a and 15b of drawing 2 and is inputted into the low pass filter of cut-off-frequency $f_s / 2$ at the interpolated signal, a sampling hold will be carried out and the signal of a smooth curve like drawing 6 (d) will be outputted. If this signal is given to the down sample means 16a and 16b, a sampling will be performed on the frequency of $m \times f_s$ ($n > m > 1$), and a signal like drawing 6 (e) will be generated. The number of data is increasing the signal shown in drawing 6 (e) by $14/11$ time with the same time interval from the signal shown in (a). In this way, the control signal of time series finer than the original sampling wave is acquired. And if the signal of this time series is given to the shift register of for example, serial in parallel out and each signal is taken out simultaneously, the control signal amended by the desired value will be acquired.

[0031] Next, this technique is collapsed, it applies to a multiplier, and the case

where more exact image normal is realized is explained. Drawing 7 shows the concept of how to shift the data to the individual difference of a convolution multiplier. Drawing 7 (a) shows the typical convolution multiplier statistically calculated from many listening people to an input sound signal. Therefore, if the image normal means 12a and 12b perform convolution using this multiplier, the remarkable listening person should be able to perceive a natural feeling of image normal. However, in many cases, a head configuration is actually large for some listening persons, or small. For this reason, by the standard multiplier, a delicate difference arises at both lugs spacing, and amendment is needed for it to each listening person.

[0032] The approach of changing the phase contrast of the sound which shifts a multiplier in time and reaches right-and-left both lugs as this amendment approach is taken. It becomes as shown in 1 tap ***** and (b) temporarily about the multiplier of drawing 7 (a). For example, it is 44.1kHz about a sampling frequency f_s . If it carries out and samples on this frequency, even if it carries out [1 tap **] a multiplier, both the lugs spacing difference is 7.7mm. It becomes a limitation. Therefore, 7.7mm It becomes impossible to correspond to the individual difference which has both the following lugs spacing. However, when an exaggerated sample is increased n times, it can respond also to the individual difference of $7.7/n$ (mm) extent.

[0033] Drawing 7 (c) carries out the exaggerated sample of the multiplier shown in (a). If it carries out [1 tap **] in this condition, a phase can be more finely shifted also by the same multiplier like drawing 7 (d) - (f). Like especially an inner phone, when both the individual delicate lugs spacing difference influences a feeling of image normal greatly, it becomes an effective approach. Thus, it can be made to orientate in the direction which wishes for an image to the listening person of arbitration.

[0034] The sound field regenerative apparatus in the 2nd example (claim 3) of this invention is explained using drawing 8 below. Drawing 8 is the block diagram showing the configuration of some sound field regenerative apparatus of this example, and shows only the circuit added to the signal-processing means [which are shown in drawing 2 / 3c and 3d] latter part. Other blocks are the same as that of the 1st example, and explanation of a whole configuration is omitted. As shown in drawing 8 , the reverberation sound creation means 17a and 17b are established to the signal-processing means [3c and 3d] output signal.

[0035] The reverberation sound creation means 17a and 17b are means to create and add the reverberation sound according to the size of sound field, and the output is given to D/A converters 4a and 4b of drawing 1 , respectively. The reverberation sound creation means 17a and 17b are constituted by carrying out

the series connection of two or more feed bottom echo circuits where time delays differ. In taking out the breadth of the space of for example, 10m order as a reverberation sound, the die length of a reverberation sound becomes 0.25 - 0.35 seconds. Moreover, when taking out the breadth of space to 10m or more sense of distance around 20m, the die length of for example, a reverberation sound is 0.7-0.9. It is set as a second. Moreover, if it is sound field playback like a large concert hall, it is comparatively long in reverberation time, and it adds so that reverberation time of low compass may be made longer than that of a loud-sound region.

[0036] Drawing 9 is the conceptual diagram of the parameter control of the sound field regenerative apparatus in the 3rd example (claim 4). The configuration of the sound field regenerative apparatus of this example is also the same as that of drawing 1 . In this example, while inputting the location of a sound source, and the size of sound field with the input means 10, theta is inputted whenever [side angle-of-reflection / which shows the distance relation between the listening person P and Image S]. For example, the value of theta will become small, if the distance from Image S becomes far as the listening person P2 of drawing 9 shows. The value of theta will become large, if distance becomes near as the listening person P1 furthermore shows. The delay and the convolution multiplier in the signal-processing means 3 shall change with the

magnitude of this side angle of reflection θ .

[0037] Drawing 10 is the block diagram of the sound field regenerative apparatus in the 4th example (claim 5) of this invention, the same part as the 1st example attaches the same sign, and those explanation is omitted. As shown in this Fig., the parameter receiving means 18 is established instead of the input means 10 of drawing 1. 18 is a means to receive the control signal for controlling a location, distance, and a feeling of breadth from the exterior, with a parameter receiving means.

[0038] With the parameter receiving means 18, reception of the control signal which included a location, distance, and the conditions of a feeling of breadth from the exterior gives the control signal to the parameter-control means 11. And the parameter-control means 11 outputs the parameter (each multiplier) which should be set up with the signal-processing means 3. The following actuation is the same as that of the 1st example mentioned above.

[0039] As mentioned above, in each example, by carrying out the exaggerated sample of the convolution multiplier for the image normal beforehand set up to the listening person of arbitration, phase adjustment can be carried out finely and it can respond. And the feeling of normal, the sense of distance, and the feeling of breadth of a sound are controllable by the control signal from the outside, and using the signal programmed beforehand, a repeat is controlled or it

becomes controllable [the feeling of normal of a sound according to the scene of an image, a sense of distance, and a feeling of breadth] combining an image etc.

[0040] In addition, although this example described the case where an input signal was a sound signal of a monophonic recording, also in the case of a stereo, it can respond by establishing an independent signal-processing means for every signal, and is effectively applied also in a two-channel loudspeaker instead of headphone or an inner phone.

[0041]

[Effect of the Invention] as mentioned above , according to this invention , by control a parameter according to the configuration of the head of a listen person , and the location of an image and a listen person and the size of magnitude , distance , or sound field which be inputted into the input means , and carry out the exaggerated sample of the convolution multiplier for make an image orientate , it can respond to both the individual delicate lugs spacing difference to the listen person of arbitration , and natural image normal and sound field playback can be realize .

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block diagram of the sound field regenerative apparatus in the 1st example of this invention.

[Drawing 2] It is the block diagram showing the configuration of the signal-processing means used for the sound field regenerative apparatus of the 1st example.

[Drawing 3] It is the explanatory view showing the principle of the image normal in the sound field regenerative apparatus of this invention.

[Drawing 4] It is the explanatory view showing the relation between a listening person and an image.

[Drawing 5] It is the graph which shows the frequency characteristics and the impulse response of both lugs.

[Drawing 6] It is the conceptual diagram showing the principle of operation from an exaggerated sample to a down sample.

[Drawing 7] It is the conceptual diagram of how to shift the convolution multiplier to a listening person's individual difference.

[Drawing 8] It is the block diagram showing the configuration of some sound field regenerative apparatus in the 2nd example of this invention.

[Drawing 9] It is the conceptual diagram of the parameter control in the sound field regenerative apparatus of the 3rd example of this invention.

[Drawing 10] It is the block diagram showing the configuration of the sound field

regenerative apparatus in the 4th example of this invention.

[Drawing 11] It is the block diagram showing the example of a configuration of the conventional image stereotaxic instrument.

[Description of Notations]

1 Signal Input Means

2, 2a, 2b A/D converter

3, 3c, 3d Signal-processing means

3a, 3b FIR filter

4a, 4b D/A converter

5a, 5b Amplifier

6a, 6b Loudspeaker

7 Listening Person

8 Signal Generation Means

9 Virtual Loudspeaker

10 Input Means

11 Parameter-Control Means

12a, 12b Image normal means

14a, 14b Exaggerated sample means

15a, 15b Filter means

16a, 16b Down sample means

17a, 17b Reverberation sound creation means

18 Parameter Receiving Means